



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

*Am*

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/671,120	09/28/2000	Eiichi Takahashi	21.1980/CJG	8624

21171 7590 06/14/2005

STAAS & HALSEY LLP  
SUITE 700  
1201 NEW YORK AVENUE, N.W.  
WASHINGTON, DC 20005

EXAMINER

SHARON, AYAL I

ART UNIT PAPER NUMBER

2123

DATE MAILED: 06/14/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

## Office Action Summary

Application No.

09/671,120

Applicant(s)

TAKAHASHI ET AL.

Examiner

Ayal I. Sharon

Art Unit

2123

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☒ Responsive to communication(s) filed on 28 February 2005.
- 2a) ☐ This action is FINAL. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 1-15 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☒ Claim(s) 8-11 is/are allowed.
- 6) ☒ Claim(s) 1,2,4-7 and 12-15 is/are rejected.
- 7) ☒ Claim(s) 3 is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 28 September 2000 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some \* c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_\_
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: \_\_\_\_\_

## **DETAILED ACTION**

### ***Introduction***

1. Claims 1-14 of U.S. Application 09/671,120 filed on 09/28/2000 are presented for examination. The application claims foreign priority to Japanese application 11-279516, filed on 09/30/1999. In the RCE filed on 2/28/2005, Applicants have amended claims 1 and 8-15.
2. In the RCE filed on 2/28/2005, Applicants have repeated the same arguments (see p.9 of the RCE) that were previously presented in their After Final Amendment filed on 2/28/05.
3. Examiner responded to those arguments in the Advisory Action dated 3/22/05.
4. In particular, Examiner notes that Applicants' amendment to the claims necessitated the new grounds for rejection in the Final Office Action. Furthermore, Applicants' further amendment to the claims in the After Final Amendment raised new issues that required further search and consideration.

### ***Allowable Subject Matter***

5. Claims 8-11 are allowed. These claims were objected in the Final Office Action dated 11/30/2004, for being dependent upon a rejected base claim, but were indicated as being allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

***Claim Objections***

6. Claim 3 is objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims. Jacobs does not expressly teach the formulas recited in Claim 3.

***Claim Rejections - 35 USC § 102***

7. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

8. The prior art used for these rejections is as follows:
9. Jacobs et al. U.S. Patent 6,385,643. (Henceforth referred to as "**Jacobs**").
- 10. Claims 1-2, 6, and 12-15 are rejected under 35 U.S.C. 102(e) as being anticipated by Jacobs.**

11. In regards to Claim 1, Jacobs teaches the following limitations:

1. A service distribution device for distributing specified services among a plurality of servers in which there is a difference in processing capacity on a network to balance the server loads, comprising:

Art Unit: 2123

Jacobs teaches (col.11, lines 9-14): "Similarly, in an embodiment, client 504, server 502, and server 503 may be both clients and servers as described above. One of ordinary skill in the art would understand that in alternate embodiments, multiple other servers and clients may be included in architecture 500 as illustrated by ellipses."

a packet capture device capturing packets transmitted through the network to calculate the server processing time and parameters to configure simulation models;

(Jacobs, especially: "D. Rerouting" at col.9, line 34 to col.10, line 5)

Jacobs teaches (col.9, lines 53-61; emphasis added): "In an embodiment, an RJVM sends a heartbeat message to other clients/servers when no other message has been sent in a pre-determined time period. **If the client/server does not receive a heartbeat message in the predetermined count time,** a failed client/server which should have sent the heartbeat, is detected. In an embodiment, a failed client/server is detected by connection timeouts or if no messages have been sent by the failed client/server in a predetermined amount of time. In still another embodiment, a failed socket indicates a failed server/client."

a server identifier recording information pertaining to the captured packets into a server log for each server;

(Jacobs, especially: "E. Multitier/Peer-to-Peer Functionality", col.10, lines 6-55)

Jacobs teaches (col.10, lines 8-27; Emphasis added): "Clustered enterprise Java™ architecture 300 supports an explicit syntax for client/server programming consistent with a multi-tier distributed processing architecture. **As an example, the following client-side code fragment writes an informational message to a server's log file:**

```
T3Client clnt=new T3Client("t3://acme:7001");
LogServices log=clnt.getT3 Service( ).log( );
log.info("Hello from a client");
```

The first line establishes a session with the acme server using the t3 protocol. If RJVMs do not already exist, each JVM constructs an RJVM for the other and an underlying TCP socket is established. The client-side representation of this session--the T3Client object--and the server-side representation communicate through these RJVMs. The server-side supports a variety of services, including database access, remote file access, workspaces, events, and logging. The second line obtains a LogServices object and the third line writes the message."

a service identifier recording information pertaining to the captured packets into a service log for each service;

(Jacobs, especially: "E. Multitier/Peer-to-Peer Functionality", col.10, lines 6-55)

Jacobs teaches (col.10, lines 8-27; Emphasis added): "The first line establishes a session with the acme server using the t3 protocol. If RJVMs do not already exist, each JVM constructs an RJVM for the other and an underlying TCP socket is established. **The client-side representation of this session--the T3Client object--and the server-side representation communicate through these RJVMs. The server-side supports a variety of services, including database access, remote file access, workspaces, events, and logging. The second line obtains a LogServices object and the third line writes the message.**"

Moreover, Jacob also teaches the following (see Fig.6a, Item 604 and associated text at col.12, lines 5-11): "In logic block 604, a naming service is contacted and an updated list of the current service providers is obtained. A getNextProvider method is called to obtain a service provider in logic block 605. Various embodiments of the getNextProvider method are illustrated in FIGS. 6b-g and described in detail below. The service is obtained in logic block 606. Failover method 508 is then called if service is not provided in logic block 606 and load balancing method 507 exits as illustrated by logic block 608. An embodiment of failover method 508 is illustrated in FIG. 7 and described in detail below."

a server modeling module setting up a simulation model for each server from the server log;  
(Jacobs, especially: col.11, lines 17-45)

Jacob teaches (See col.11, lines 17-27): "**RA RMI stub 580 is a Smart stub which is able to find out about all of the service providers and switch between them based on a load balancing method 507 and/or failover method 508.** In an embodiment, an RA stub 580 includes a replica handler 506 that selects an appropriate load balancing method 507 and/or failover method 507. In an alternate embodiment, a single load balancing method and/or single failover method is implemented. In alternate embodiments, replica handler 506 may include multiple load balancing methods and/or multiple failover methods and combinations thereof."

Jacob teaches (See col.11, lines 41-43): "**Load balance method 507 may switch to server 503 because server 502 is overloaded with service requests.**"

a service modeling module setting up a simulation model for each service from the service log;  
(Jacobs, especially: Fig. 6a, Item 604 and associated text in col.11, line 52 to col. 12, line 15)

a simulator reading in the server model and the service model and running each simulation; and

(Jacobs, especially: Fig.5b, Item 504, and associated text at col.11, lines 1-17)

a server selection module selecting and specifying an optimum server to distribute services to based on a simulator result.

(Jacobs, especially: See col.11, lines 37-52)

Jacob teaches (See col.11, lines 41-43): **"Load balance method 507 may switch to server 503 because server 502 is overloaded with service requests."**

12. In regards to Claim 2, Jacobs teaches the following limitations:

2. The service distribution device of claim 1, further comprising a packet relay device obtaining packets using a packet capture module mounted on said packet relay device, which relays packets between a client and the servers.

(Jacobs, especially: Fig.5b, Items 501-4 and associated text.)

Examiner finds it to be inherent that that a communication medium such as the one taught in Jacobs (Fig.5b, Item 501) will relay packets between a client (Fig.5b, Item 504) and the servers (Fig.5b, Items 502-3), otherwise there will be no communication between the client and the servers.

13. In regards to Claim 6, while Jacobs teaches the following:

6. The service distribution device of claim 1, wherein said server selection module determines a standard value using an output of a single simulation run for each service by said simulator, and determines that a high load state exists if a difference between, or the ratio of, the standard value and the output of the simulation of a plurality of sessions exceeds a predetermined threshold.

(Jacobs, especially: See "D. Rerouting" at col.9, line 34 to col.10, line 5; and col.11, lines 37-52)

Jacob teaches (See col.11, lines 41-43): **"Load balance method 507 may switch to server 503 because server 502 is overloaded with service requests."**

Jacobs teaches (col.9, lines 53-61; emphasis added): "In an embodiment, an RJVM sends a heartbeat message to other clients/servers when no other message has been sent in a pre-determined time period. **If the client/server does not receive a heartbeat message in the predetermined count time**, a failed client/server which should have sent the heartbeat, is detected. In an embodiment, a failed client/server is detected by connection timeouts or if no messages have been sent by the failed client/server in a predetermined amount of time. In still another embodiment, a failed socket indicates a failed server/client."

14. In regards to Claim 12, Jacobs teaches the following limitations:

12. A service distribution device for distributing specified services among a plurality of servers in which there is a difference in processing capacity to balance server loads, comprising:

a server modeling module generating a simulation model for each server and a service modeling module generating a simulation model for each service based on a server log and a service log of captured server communications;

Jacobs teaches (col.10, lines 8-27; Emphasis added): "The first line establishes a session with the acme server using the t3 protocol. If RJVMs do not already exist, each JVM constructs an RJVM for the other and an underlying TCP socket is established. **The client-side representation of this session--the T3Client object--and the server-side representation communicate through these RJVMs. The server-side supports a variety of services, including database access, remote file access, workspaces, events, and logging.** The second line obtains a LogServices object and the third line writes the message."

a simulator reading the server models and the service models and running a plurality of simulations; and

Jacobs teaches (col.11, lines 9-14): "Similarly, in an embodiment, client 504, server 502, and server 503 may be both clients and servers as described above. One of ordinary skill in the art would understand that in alternate embodiments, multiple other servers and clients may be included in architecture 500 as illustrated by ellipses."

a server selection module determining which servers have low loads based on results of the simulations and selecting the servers with low loads to receive the services.

Jacob teaches (See col.11, lines 41-43): "**Load balance method 507 may switch to server 503 because server 502 is overloaded with service requests.**"

15. In regards to Claim 13, Jacobs teaches the following limitations:

13. A method for distributing specified services among a plurality of servers in which there is a difference in processing capacity to balance server loads, comprising:

generating a simulation model for each server and each service based on a server log and a service log of captured server communications;

Jacobs teaches (col.10, lines 8-27; Emphasis added): "The first line establishes a session with the acme server using the t3 protocol. If RJVMs



Art Unit: 2123

do not already exist, each JVM constructs an RJVM for the other and an underlying TCP socket is established. **The client-side representation of this session--the T3Client object--and the server-side representation communicate through these RJVMs. The server-side supports a variety of services, including database access, remote file access, workspaces, events, and logging.** The second line obtains a LogServices object and the third line writes the message.”

running a plurality of simulations using the server and service models; and

Jacobs teaches (col.11, lines 9-14): “Similarly, in an embodiment, client 504, server 502, and server 503 may be both clients and servers as described above. One of ordinary skill in the art would understand that in alternate embodiments, multiple other servers and clients may be included in architecture 500 as illustrated by ellipses.”

determining which servers have low loads based on results of the simulations and selecting the servers with low loads to receive the services.

Jacob teaches (See col.11, lines 41-43): “**Load balance method 507 may switch to server 503 because server 502 is overloaded with service requests.**”

16. In regards to Claim 14, Jacobs teaches the following limitations:

14. A computer-readable storage controlling a computer to distribute services among a plurality of servers in which there is a difference in processing capacity and comprising a process of:

generating a simulation model for each server and each service based on a server log and a service log of captured server communications;

Jacobs teaches (col.10, lines 8-27; Emphasis added): “The first line establishes a session with the acme server using the t3 protocol. If RJVMs do not already exist, each JVM constructs an RJVM for the other and an underlying TCP socket is established. **The client-side representation of this session--the T3Client object--and the server-side representation communicate through these RJVMs. The server-side supports a variety of services, including database access, remote file access, workspaces, events, and logging.** The second line obtains a LogServices object and the third line writes the message.”

running a plurality of simulations using the server and service models; and

Jacobs teaches (col.11, lines 9-14): “Similarly, in an embodiment, client 504, server 502, and server 503 may be both clients and servers as described above. One of ordinary skill in the art would understand that in alternate

Art Unit: 2123

embodiments, multiple other servers and clients may be included in architecture 500 as illustrated by ellipses."

determining which servers have low loads based on results of the simulations and selecting the servers with low loads to receive the services.

Jacob teaches (See col.11, lines 41-43): **"Load balance method 507 may switch to server 503 because server 502 is overloaded with service requests."**

17. In regards to Claim 15, Jacobs teaches the following limitations:

15. (new) A method for distributing services among a plurality of servers in which there is a difference in processing capacity, comprising:

capturing network communication of at least one of the servers;  
adding information about the captured network communication to a log;  
simulating the servers based on the log;

Jacobs teaches (col.10, lines 8-27; Emphasis added): "The first line establishes a session with the acme server using the t3 protocol. If RJVMs do not already exist, each JVM constructs an RJVM for the other and an underlying TCP socket is established. **The client-side representation of this session--the T3Client object--and the server-side representation communicate through these RJVMs. The server-side supports a variety of services, including database access, remote file access, workspaces, events, and logging.** The second line obtains a LogServices object and the third line writes the message."

determining at least one of the servers having a low load based on the simulation; and distributing a service to the determined at least one of the servers.

Jacob teaches (See col.11, lines 41-43): **"Load balance method 507 may switch to server 503 because server 502 is overloaded with service requests."**

### ***Claim Rejections - 35 USC § 103***

18. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the

Art Unit: 2123

invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

19. The prior art used for these rejections is as follows:

20. Jacobs et al. U.S. Patent 6,385,643. (Henceforth referred to as "**Jacobs**").

21. Jain, R. The Art of Computer Systems Performance Analysis. © 1991. pp.624-

626. Specifically, the section titled "Symbols Frequently Used in Queueing Analysis." (Henceforth referred to as "**Jain**").

22. Kleinrock, L. "On the Modeling and Analysis of Computer Networks." Proc. of the IEEE. Aug.1993. pp.1179-1191. (Henceforth referred to as "**Kleinrock**").

23. Zhu, H. "Adaptive Load Sharing for Clustered Digital Library Services". The 7<sup>th</sup> Int'l Symposium on High Performance Computing. July 31, 1998. pp. 235-242. (Henceforth referred to as "**Zhu**").

24. The claim rejections are hereby summarized for Applicant's convenience. The detailed rejections follow.

**25. Claim 4 is rejected under 35 U.S.C. 103(a) as being unpatentable over Jacobs in view of Jain and further in view of Kleinrock.**

26. In regards to Claim 4, Jacobs does not expressly teach the following limitations:

4. The server distribution device of claim 1, wherein said service modeling module calculates the following parameters from the service log by constructing a service model for each service:

a ratio of the number of sessions for each service to the number of sessions for all services,

a session starting frequency or time interval,

a number of transmissions between the client and server per session,

a client response size, packet size, and packet count per transmission,

a server response size, packet size, and packet count per transmission, and

a time from the server response until the client response.

Jain, on the other hand, teaches the following parameters which correspond to those claimed by the Applicants:

- The ratio of the parameters  $D_i$  and  $D$ , (Total service demand on server 'i', and Total service demand on all servers), as taught by Jain, produces a ratio that corresponds to Applicants' claimed limitation:

a ratio of the number of sessions for each service to the number of sessions for all services,

- The parameter  $\tau$ , (Inter-arrival time), as taught by Jain, corresponds to Applicant's claimed limitation:

a session starting frequency or time interval,

- The parameter  $V_i$ , (Number of visits to service center i), as taught by Jain, corresponds to Applicant's claimed limitation:

a number of transmissions between the client and server per session,

- The parameter  $I$ , (Idle time duration for a server), as taught by Jain, corresponds to Applicant's claimed limitation:

a time from the server response until the client response.

It would have been obvious to one of ordinary skill in the art to modify the teachings of Jacobs with those of Jain, because Jain's symbols are "Symbols Frequently Used in Queueing Analysis." (This is the title that Jain gave to the section of his article which discusses these parameters).

Kleinrock teaches the following parameters which correspond to those claimed by the Applicants:

a client response size, packet size, and packet count per transmission,

a server response size, packet size, and packet count per transmission, and

Kleinrock teaches (p.1180, col.2, para.3) that "The modification is to assume that message lengths are all the same (rather than the exponential assumption above), and that the topology is a tandem network.

It would have obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Jacobs with those of Kleinrock, Because Jacobs expressly teaches (col.5, lines 27-32) that "Both the server parameters and required service characteristics are inputs to modeling process such as is described in [two other Kleinrock references related to modeling queuing computer networks]."

**27. Claim 5 is rejected under 35 U.S.C. 103(a) as being unpatentable over Jacobs in view of Kleinrock.**

28. In regards to Claim 5, Jacobs does not expressly teach the following:

5. The service distribution device of claim 1, wherein said simulator performs a simulation using the server model and the service model and generates a mean value or a median value of a session time for the specific service.

Kleinrock, on the other hand, does expressly teach:

a) generating the mean delay time of a system "One of the first general results was an exact expression for the mean delay experienced by a message as it passed through a network ..." (See p.1179, col.2, paragraph 4).

b) generating the mean response time of a system "In addition, we let  $T_i$  be the mean response time of this little queueing system." (See p.1180, col.2, Eq.4)

It would have obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Jacobs with those of Kleinrock, Because Jacobs expressly teaches (col.5, lines 27-32) that "Both the server parameters and required service characteristics are inputs to modeling process such as is described in [two other Kleinrock references related to modeling queuing computer networks]."

**29. Claim 7 are rejected under 35 U.S.C. 103(a) as being unpatentable over Jacobs in view of Zhu.** 

30. In regards to Claim 7, while Jacobs teaches a "... round-robin scheduling balances the load among the servers" (Jacobs: col.7, lines 8-9), Jacobs does not expressly teach the following:

7. The service distribution device of claim 6, wherein when said server selection module receives a server distribution query, said server selection module sets a server permission to be a starting frequency of the session that will cause a high load state for the service in question for each server, and specifies a server having the biggest difference between the session starting frequency and the permission as a server for distribution.

Zhu does teach these limitations (see Section 3.2, "Policies for Node Selection and Load Collection").

It would be obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Jacobs with those of Zhu, because doing so enables "... each processor to make a decision in a distributed manner and choose the best server node for redirection ..." (Zhu, Section 3.2, first paragraph).

***Response to Amendment***

***Re: Claim Rejections - 35 USC § 102***

31. Examiner has found applicants' arguments regarding the Caswell reference to be persuasive (see RCE filed 2/28/2005, p.10, paragraph 5) and has withdrawn the reference.
32. New art rejections have been applied.

***Re: Claim Rejections - 35 USC § 103***

33. Examiner has found applicants' arguments regarding the Caswell reference to be persuasive (see RCE filed 2/28/2005, p.10, paragraph 5) and has withdrawn the reference.
34. New art rejections have been applied.

***Correspondence Information***

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Ayal I. Sharon whose telephone number is (571) 272-3714. The examiner can normally be reached on Monday through Thursday, and the first Friday of a biweek, 8:30 am – 5:30 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Leo Picard can be reached at (571) 272-3749.

Any response to this office action should be faxed to (703) 872-9306, or mailed to:

Art Unit: 2123

USPTO  
P.O. Box 1450  
Alexandria, VA 22313-1450

or hand carried to:


USPTO  
Customer Service Window  
Randolph Building  
401 Dulany Street  
Alexandria, VA 22314

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the Tech Center 2100 Receptionist, whose telephone number is (571) 272-2100.

Ayal I. Sharon

Art Unit 2123

June 8, 2005

  
W. Turner  
Primary Examiner  
TC 2100  
AU 2123